

Neurological Disease Detection and Monitoring from Voice Production

Pedro Gómez-Vilda¹, Victoria Rodellar-Biarge¹, Víctor Nieto-Lluis¹,
Cristina Muñoz-Mulas¹, Luis Miguel Mazaira-Fernández¹, Carlos Ramírez-Calvo²,
Mario Fernández-Fernández², and Elvira Toribio-Díaz²

¹ Grupo de Informática Aplicada al Tratamiento de Señal e Imagen, Facultad de Informática,
Universidad Politécnica de Madrid, Campus de Montegancedo, s/n, 28660 Madrid
pedro@pino.datsi.fi.upm.es

² ENT and Neurology Services, Hospital del Henares, Avda. Marie Curie s/n,
28822 Coslada, Madrid, Spain

Abstract. The dramatic impact of neurological degenerative pathologies in life quality is a growing concern. It is well known that many neurological diseases leave a fingerprint in voice and speech production. Many techniques have been designed for the detection, diagnose and monitoring the neurological disease. Most of them are costly or difficult to extend to primary attention medical services. Through the present paper it will be shown how some neurological diseases can be traced at the level of phonation. The detection procedure would be based on a simple voice test. The availability of advanced tools and methodologies to monitor the organic pathology of voice would facilitate the implantation of these tests. The paper hypothesizes that some of the underlying mechanisms affecting the production of voice produce measurable correlates in vocal fold biomechanics. A general description of the methodological foundations for the voice analysis system which can estimate correlates to the neurological disease is shown. Some study cases will be presented to illustrate the possibilities of the methodology to monitor neurological diseases by voice.

Keywords: neurological disease, voice production, pathology grading and monitoring, spasmodic dysphonia, Parkinson Disease.

1 Introduction

Voice resulting from altered phonation due to neurological causes may be a most valuable report of the etiology and progress of neural diseases affecting the production of voice, such as pathologies resulting in voice tremor [1]. These would include some kinds of spasmodic dysphonia, stammering and Parkinson. The possibility of early detection in the first stages of Parkinson's Disease (PD) may grant a better preventive treatment reducing the progress of illness [2]. Monitoring treatment by objective methods is also an important goal, especially in modifying or defining new protocols. The deepest foundations of the methodology proposed in this paper are to be found in tracking the malfunctioning of neurological and neuromuscular paths involved in voice production (see Fig. 1).

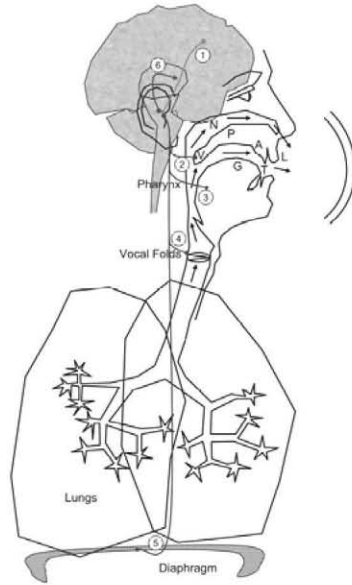


Fig. 1. Simplified view of main neural pathways involved in the production of phonation: 1. Links from linguistic neuromotor cortex to Basal Ganglion relay stages. 2. Branch of the X nerve acting on the velo-pharyngeal switch. 3. Idem acting on the retro-lingual switch connected to the epiglottal switch. 4. Branch of the laryngeal nerve acting on the transversal and oblique arytenoid and cricothyroid muscles responsible for the vocal fold adduction and abduction. 5. Branch of the vagus nerve (phrenic) actuating on the diaphragmatic muscles. 6. Feedback loop in Basal Ganglia damping muscular tone (its malfunction is involved in PD).

These comprise links from the neuromotor linguistic cortex [3] to the subthalamic region [4] and through the laryngeal nerve and their associated pathways [5][6] to the muscles activating the thyro-arytenoid structure, responsible in the last term of vocal fold stretching, adduction and abduction (Superior Laryngeal Nerve, Internal and External Laryngeal Branches of the Inferior Laryngeal Nerve, Transverse and Oblique Arytenoid Muscles, and Cricothyroid Muscles). Any alteration in the functionality of these pathways and in the associated muscles will result in temporary distortions of the stiffness estimates of the vocal folds biomechanics. The aim of this paper is to give some phenomenological account in detecting and grading the neurological disease using biomechanical correlates obtained from the inverse filtering of voice.

2 Estimating Neurological Correlates from Biomechanics

Inverse filtering and modeling can be used to estimate the rough biomechanical parameters of a simple 2-mass model of the vocal folds as shown in [7]. A separate estimate may be produced for each glottal cycle (epoch). Therefore a series of around 50-100 estimates for each parameter of dynamic mass, losses and stiffness of the vocal fold body could be drawn from a segment of voice lasting 0.5 seconds,

depending on the speaker's gender. As tremor in voice resulting from neurological disease may contain up to 6-8 spasms or vibrato cycles in one second, segments shorter than 0.5 s would not be recommendable for this kind of analysis [8]. The working hypothesis would then be that tremor may leave correlates in different biomechanical as well as distortion parameters. Specifically, it is hypothesized that the influence of the neurological disease has to leave a mark in the tension estimated on the vocal folds as a cyclic alteration from stiffness estimates. The question now is how to estimate cyclic behavior in stiffness cycle by cycle. A simple and feasible approach may be AR modeling by adaptive inverse filtering. Specifically ξ_n being the stiffness estimate at phonation cycle n , its AR model would be described as:

$$\xi_n = \sum_{i=1}^K a_i \xi_{n-i} + \varepsilon_n \quad (1)$$

where $\mathbf{a}=\{a_i\}$ are the model coefficients. The estimation of the coefficients is feasible by the minimization of the modeling error ε_n in terms of Least Mean Squares (LMS). A possible way to implement such estimation is by means of adaptive lattice filters [9]. An adaptive lattice filter may be defined as an operator $\Phi_{Kn}[\cdot]$ which may produce an output error $\varepsilon_K(n)$ minimum in terms of LMS for a given time window W_K which may slide along the time index n rendering a sequence of sub-optimal models from a non-stationary input series under a certain adaptation factor β :

$$\{\varepsilon_{Kn}, \mathbf{c}_{Kn}\} = \Phi_{Kn}\{\xi_n, W_K, \beta\} \quad (2)$$

Once the lattice adaptive model has been fitted to the input series, either its pivoting coefficients \mathbf{c}_{Kn} or those of the equivalent transversal model \mathbf{a}_{Kn} may be used as model and series descriptors. Both sets of coefficients are related by Levinson-Durbin iteration:

$$\mathbf{a}_{kn} = \mathbf{a}_{k-1n} - c_{kn} \tilde{\mathbf{a}}_{k-1n} \quad (3)$$

where $\tilde{\mathbf{a}}$ is the order-reversing operation on vector \mathbf{a} . In the present study pivoting coefficients will be preferred, as they are pre-normalized in the interval (-1, 1), thus being described in range which allows easy comparison among different results. In the present case the three lowest-order pivoting coefficients $\{c_{1n}, c_{2n}, c_{3n}\}$ will be used as descriptors of the stiffness cyclic characteristics, as shown in the sequel.

3 Materials and Methods

A preliminary study was conducted to envision the possibilities of the methodology proposed in section 2. This study has a phenomenologic descriptive character, and is based in four specific cases of female voice drawn from a database being recorded currently by the ENT and Neurology Services in Hospital del Henares. They correspond to a 34-year old normophonic subject (NF, case #346), a 22-year old subject affected by a Left Vocal-Fold Cyst (LVFC, case #341), a 45-year old subject affected by Spasmodic Dysphonia (SD, case #308), and a 72-year old subject affected by Parkinson Disease (PD, case #337523). The tests included ORL inspection by endoscopy to assess organic pathological condition and GRBAS evaluation. Subjects

NF and LVFC corresponding respectively to a normophonic and a dysphonic case with clear organic pathology were included in the study to serve as controls in contrast with subjects with clear tremor pathology (SD and PD). Voice and speech tests were recorded including the five cardinal vowels in Spanish [a, e, i, o, u], target words to measure the velo-pharyngeal switch and short sentences in which these words appeared in co-articulation. For the present study segments of vowel [a] 0.5 s long were used. The protocol carried out was the following:

- Inverse filtering was used to estimate the glottal source as mentioned before.
- The biomechanical parameters corresponding to vocal fold body mass, stiffness and losses were estimated as by [8].
- The stiffness parameter was unbiased and smoothened using low-pass filtering.
- The cyclic descriptors $\{c_{1n}, c_{2n}, c_{3n}\}$ were estimated using adaptive inverse modeling as by (2)-(3).

The glottal source reconstructed for each case in detail and in its full extension are given in Fig. 2. The glottal arch (Fant's LF-pattern [10]) is given in the upper part of each template in blue, the associated glottal flow given in green.

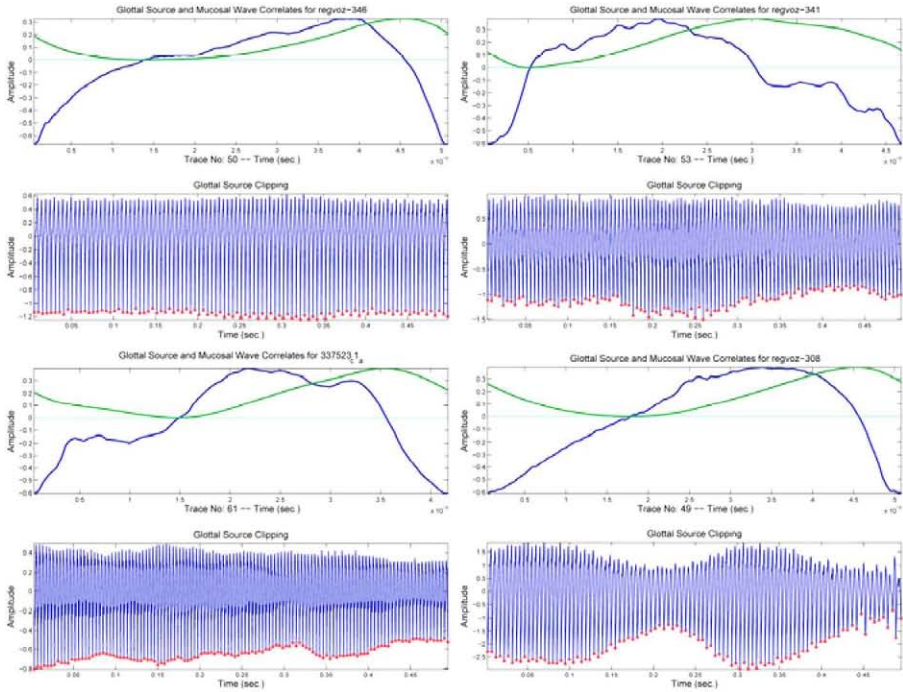


Fig. 2. Glottal source estimates of the cases studied (upper templates - one glottal cycle detail; lower templates - full series). Top left: Normophonic female voice (#346). Top right: LVFC female voice (#341). Bottom left: PD female voice (#337523). Bottom right: SD female voice (#308).

It may be seen that the most regular pattern accordingly with Fant's LF-pattern corresponds to the uppermost left (NF), the case LVFC being the most irregular accordingly with phonation regularity rules. Cases PD and SD cannot be considered pathological under the organic point of view, although the glottal cycle shown for SD corresponds to an epoch with marked dystonia in the middle of the recording, and is thus affected by a low tilt slope in the initial part of the cycle. The main difference between SD and PD is that the tremor is more evident in SD, as the spasms reduce sensibly the amplitude of the glottal source and the phonation. Cyclic behavior is also different in both cases, as this study is intended to show.

4 Results and Discussion

The results drawn from the analysis of the four cases mentioned are given in Fig. 3 to Fig. 6. The first case corresponds to the normophonic (NF) case. It is worth mentioning the stability in the stiffness estimate along 96 epochs, the 1-3 inter-quartile distance being 1.53% the median value. This is not the case for the SD, which shows a ratio around 75%, or that of LVFC with around 20%, whereas the PD case shows only an 8% for this ratio.

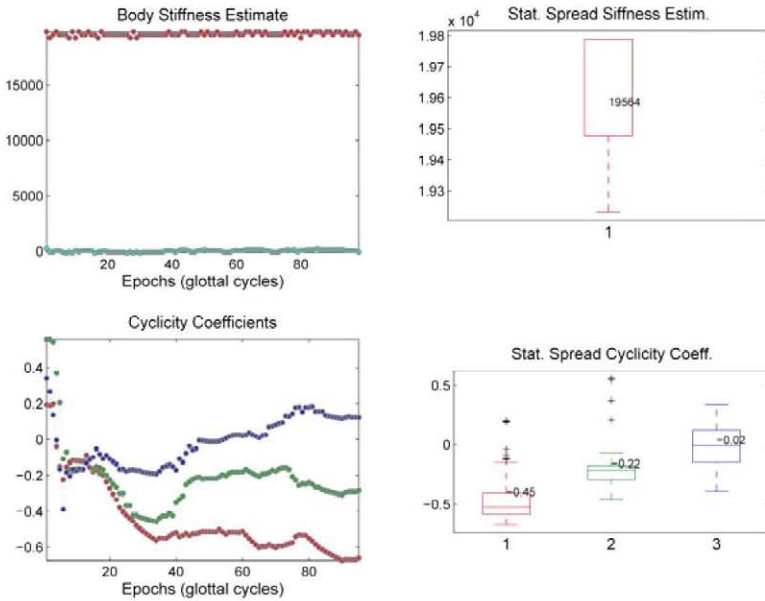


Fig. 3. Cyclic analysis of normophonic (NF) female voice (#346). Top left: Rough Body Stiffness Estimate (red) and its smoothen unbiased correlate (blue). Top right: Statistical Dispersion of the Rough Estimate. Bottom left: First, second and third order Cyclic Coefficients from adaptive estimation (respectively in red, green and blue). Bottom right: Statistical dispersion of the Cyclic Coefficients. The Rough Stiffness Estimate shows high stability and low statistical dispersion.

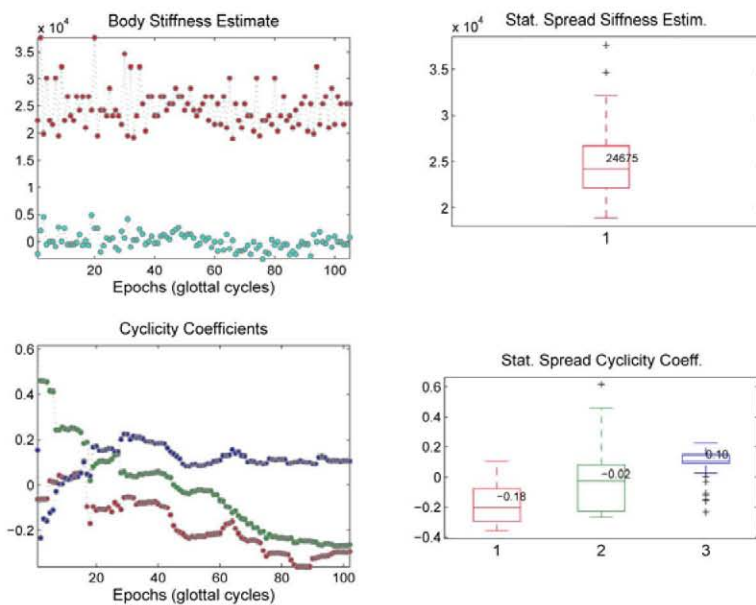


Fig. 4. Cyclic analysis of LVFC female voice (#341). See the more disordered pattern in the body stiffness estimate (top left) showing larger tension and statistical dispersion.

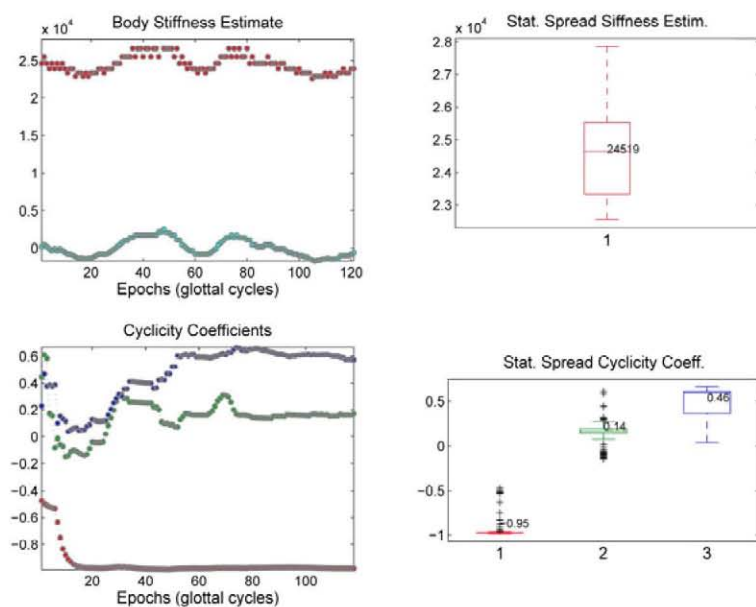


Fig. 5. Cyclic analysis of PD female voice (#337523). See that the body stiffness estimate (top left) shows a cyclic behavior in this case.

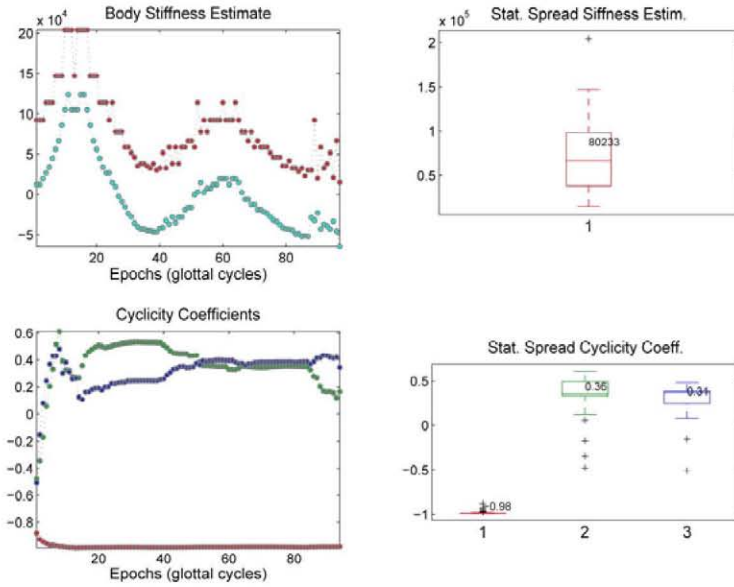


Fig. 6. Cyclic analysis of SD female voice (#308). See that the body stiffness estimate (top left) shows stronger oscillations and dispersion associated to hyper- and hypo-tonic epochs.

In all cases the adaptive estimation shows some instability during the first epochs, which converges at the end of the interval. To give a global value for each of the cyclic parameters, their medians have been estimated and listed in Table 1.

Table 1. Parameter values for the four cases studied

Subject/Parameter	Pitch (Hz)	C_1	C_2	C_3	Stiffness (g.s^{-2})
#346 (34y NF)	199	-0.45	-0.22	-0.02	19,564
#341 (22y LVFC)	215	-0.18	-0.02	0.10	24,675
#308 (45y SD)	199	-0.98	0.36	0.31	80,233
#337523 (72y PD)	248	-0.95	0.14	0.46	24,519

The first cyclic coefficient (C_1) is relatively low for NF and LVFC corresponding to normal and non-neurological diseases, as compared with the two tremor cases (SD and PD), where it is near its lowest limit (-1). Although tremor is much more evident and stronger in SD than in PD, this coefficient is similar in both cases, marking the presence of oscillation. Coefficient C_2 is negative or near zero for NF and LVFC, and clearly positive in SD and PD. C_3 it is almost zero for NF and larger in SD and PD, marking the pathological condition independently from its neurological etiology.

5 Conclusions

Some conclusions may be drawn from what has been exposed. The first one is that the main working hypothesis is verified: tremor pathology appears in the stiffness

biomechanical parameter of the vocal fold body, and is clearly observable in the neurological disease cases contrasted with the healthy or non-neurological disease ones. The second conclusion is that neurological disease is observed both in PD and SD as an oscillation in the stiffness estimate, as expected, eventually being much stronger in the second case. The third conclusion is that the cyclic behavior observed in the stiffness may be brought to global estimates using an order-3 adaptive AR model based on the pivoting coefficients of an adaptive lattice filter. The fourth one is that the three cyclic estimates derived from such model do show a relationship to the pathological/non-pathological condition, and in the first case, between the larger to the smaller tremor amplitude. Obviously the material used in the present study is of limited extent and does not permit to generalize these observations, but at least may signal a possible way for further study when a larger database would be available. If confirmed, the described methodology could open the way to new monitoring protocols for the neurologic disease of easy and simple use and at a low cost.

Acknowledgments. This work has been funded by grants TEC2006-12887-C02-01/02 and TEC2009-14123-C04-03 from Plan Nacional de I+D+i, Ministry of Science and Technology.

References

1. Pantazis, Y., Koutsogiannaki, M., Stylianou, Y.: A novel method for the extraction of tremor. In: Proc. of MAVEBA 2007, pp. 107–110. Florence University Press (2007)
2. Das, R.: A comparison of multiple classification methods for diagnosis of Parkinson disease. *Expert Systems with Applications* 37, 1568–1572 (2010)
3. Rauschecker, J.P., Scott, S.K.: Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience* 12, 718–724 (2009)
4. Törnqvist, A.L., Schalén, L., Rehncrona, S.: Effects of different electrical parameter settings on the intelligibility of speech in patients with Parkinson's Disease treated with subthalamic deep brain stimulation. *Mov. Disord.* 20, 416–423 (2004)
5. Eckley, C.A., Sataloff, R.T., Hawkshaw, M., Spiegel, J.R., Mandel, S.: Voice range in superior laryngeal nerve paresis and paralysis. *J. Voice* 12, 340–348 (1998)
6. Luschei, E.S., Ramig, L.O., Baker, K.L., Smith, M.E.: Discharge characteristics of laryngeal single motor units during phonation in young and older adults and in persons with Parkinson disease. *J. Neurophysiol.* 81, 2131–2139 (1999)
7. Gómez, P., Fernández, R., Rodellar, V., Nieto, V., Álvarez, A., Mazaira, L.M., Martínez, R., Godino, J.I.: Glottal Source Biometrical Signature for Voice Pathology Detection. *Speech Communication* 51, 759–781 (2009)
8. Tsanas, A., Little, M.A., McSharry, P.E., Ramig, L.O.: Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests. *IEEE Trans. Biomed. Eng.* 57, 884–893 (2009)
9. Deller, J.R., Proakis, J.G., Hansen, J.H.L.: *Discrete-Time Processing of Speech Signals*. Macmillan Pub. Co., Englewood Cliffs (1993)
10. Fant, G., Liljencrants, J.: A four-parameter model of glottal flow. *KTH STL-QPSR* 26(4), 1–13 (1985)